



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2019

**Do psychological defeaters undermine foundationalism in moral
epistemology? : a critique of Sinnott-Armstrong's argument against ethical
intuitionism**

Schwind, Philipp

DOI: <https://doi.org/10.1007/s10677-019-10023-9>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-176818>

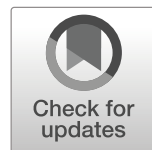
Journal Article

Published Version

Originally published at:

Schwind, Philipp (2019). Do psychological defeaters undermine foundationalism in moral epistemology?
: a critique of Sinnott-Armstrong's argument against ethical intuitionism. *Ethical Theory and Moral
Practice*, 22(4):941-952.

DOI: <https://doi.org/10.1007/s10677-019-10023-9>



Do Psychological Defeaters Undermine Foundationalism in Moral Epistemology? - a Critique of Sinnott-Armstrong's Argument against Ethical Intuitionism

Philipp Schwind¹ 

Accepted: 7 August 2019 / Published online: 14 August 2019

© Springer Nature B.V. 2019

Abstract

Foundationalism in moral epistemology is a core tenet of ethical intuitionism. According to foundationalism, some moral beliefs (such as Ross' list of prima facie duties) can be known without inferential justification; instead, all that is required is a proper understanding of the beliefs in question. In an influential criticism against this view, Walter Sinnott-Armstrong has argued that certain psychological facts undermine the reliability of moral intuitions. He claims that foundationalists would have to show that non-inferentially justified beliefs are not subject to those defeaters, but this would already constitute a form of inference and hence undermine the possibility of noninferential justification. The goal of my paper is to defend foundationalism against Sinnott-Armstrong's criticism. After presenting his challenge, I first argue that the most promising objection to it fails. This objection makes the case that defeater-defeaters are not part of the justification, but merely preserve the justification which the original claim provides. I object to this argument by distinguishing between weak and robust defeaters; only weak defeaters, I argue, fall outside the scope of justification, and it is an open question whether Sinnott-Armstrong's defeaters fall into that category. This leads the way to my own criticism of Sinnott-Armstrong's challenge: foundationalists in moral epistemology are entitled to the use of defeater-defeaters as part of the justification for moral beliefs as long as those defeater-defeaters themselves do not entail moral claims. Therefore, Sinnott-Armstrong's challenge does not undermine foundationalism.

Keywords Ethical intuitionism · Foundationalism · Justification · Moral epistemology · Experimental philosophy · Defeater

1 Introduction

One of the main reasons for the revival of ethical intuitionism over the last twenty years or so has been the development of epistemological theories which preserve the core claim of intuitionism that at least some moral beliefs can be non-inferentially justified, but entail fewer of the

✉ Philipp Schwind
philipp.schwind@philos.uzh.ch

controversial philosophical commitments which contributed to the decline of earlier intuitionism. Three respects in which contemporary intuitionists such as Robert Audi have adopted a weaker position than their predecessors are: First, an agent can possess self-evident moral knowledge even if she is not aware that her beliefs are self-evident; second, to acquire the knowledge that a proposition is self-evident, agents are allowed to make inferences as long as they merely clarify the meaning of the proposition in question and do not rely on external premises; and third, even though self-evidence is sufficient for justifying a proposition, this does not exclude the possibility of additional forms of justification for the same proposition (see Audi 2004: 40–54).

Despite these concessions, “moderate intuitionism” has failed to convince many of its critics. This is because its core claim, foundationalism in epistemology, is often also considered its core problem.¹ According to foundationalism, there are moral propositions which are such that in virtue of understanding them, “one is justified in believing the proposition [...]—this explains why such a truth is evident *in itself*. Second, if one believes the proposition on the *basis* of that understanding, then one knows it.”²

What makes foundationalism attractive is that it offers a solution to the problem of how moral beliefs can be justified. If a belief’s justification depends on other beliefs which also stand in need of justification, there looms the threat of an infinite regress or a justificatory circle.³ In response to this problem, non-inferential justification functions as a regress stopper. According to intuitionists, fundamental moral obligations such as the *prima facie* duties of veracity, benevolence or reparation serve as fixed points in our normative framework which do not stand in need of further justification and from which our duties in a concrete situation can be derived.

Many have taken issue with this claim. There are two ways in which critics have attacked the possibility of non-inferential justification. On conceptual grounds, the case can be made that the idea of foundationalism is incoherent or stands in tension with other plausible beliefs. Alternatively, many attack foundationalism on empirical grounds, with arguments to the effect that certain psychological facts undermine the reliability of our intuitions, and this makes non-inferential justification impossible.⁴ Over the past ten years, Walter Sinnott-Armstrong has put forward a version of the second kind of criticism in a series of publications.⁵

The goal of this paper is to defend foundationalism against Sinnott-Armstrong’s criticism. After presenting his argument (section 2), I argue that the most promising objection to it fails (sections 3 and 4). This leads the way to my own criticism (section 5), which should not only

¹ Besides foundationalism in epistemology, a second core claim of ethical intuitionism is non-naturalist realism (an exception is A. C. Ewing, who holds the first and third core claim of intuitionism, but rejects in his later work moral realism). Third, many but not all intuitionists defend ethical pluralism. While classical intuitionism is usually understood as the combination of these three claims, they can be held independently.

² Audi (2004: 48f.); see also Audi (1999: 206) and (2008). For Ross’s conception of self-evidence, see Ross (1930: 20 fn. 1 and 29f). Other defenders of foundationalist moral epistemology include Shafer-Landau (2003: 247) and Stratton-Lake (2002: 18–23). It has been argued, e.g. by Huemer (2005: 106), that intuitions are intellectual seemings, not self-evident beliefs. In what follows, I will limit my discussion to self-evidence accounts since this is the explicit target of Sinnott-Armstrong’s challenge to intuitionism.

³ See Sinnott-Armstrong (2006b: 74–77). The problem of course extends beyond morality, but the focus of this paper is on the case of moral beliefs.

⁴ Insofar as any ethical theory needs to account for the epistemic justification of moral beliefs, the problem raised by Sinnott-Armstrong is not limited to intuitionism: if an epistemic regress, circular forms of justification and foundationalism are the only options for solving this task, the claim that foundationalism lacks plausibility limits the options for all ethical theories. This is, however, especially problematic for intuitionism, because intuitionism—as traditionally understood—is tied to foundationalism.

⁵ See Sinnott-Armstrong (2002), (2006a), (2006b: ch. 9), (2008) and (2011). Unless indicated otherwise, I will focus on the latest formulation of his argument.

result in a better understanding of the epistemology that intuitionism is committed to, but also contribute to the debate about the nature of defeaters and defeater-defeaters.

2 Sinnott-Armstrong's Argument against Foundationalism

In line with Ross and Audi, Sinnott-Armstrong takes foundationalism in moral epistemology to hold that moral agents can be justified in believing at least some moral beliefs non-inferentially, i.e. “regardless of whether the believer is committed to any justificatory inferential structure” which he defines as “a set of propositions where some propositions provide epistemic support for others” (Sinnott-Armstrong 2011: 13). The claim is not merely that the agent need not *actually* go through an inference (even if she is in possession of a justificatory inferential structure for the moral belief in question), but the stronger requirement that an agent can be justified even without having a *disposition* to accept a justificatory inferential structure.

Sinnott-Armstrong questions this assumption. He does not do so by arguing that non-inferential justification is an incoherent or misconceived idea (the first strategy mentioned above). Instead, he takes an indirect route: he attempts to show that empirical evidence against the reliability of moral judgments *in general* makes it impossible to claim that a subgroup of these judgments is foundationally justified. That is, even if non-inferential justification turns out to be a viable epistemic tool for other kinds of beliefs, the specific features of moral belief formation undermine the applicability of such justification in the moral realm.

Here is Sinnott-Armstrong's case for the empirically informed premise of his argument: A wide range of experiments from empirical moral psychology leave little doubt that people's moral judgments are affected by non-truth-tracking factors such as implicit biases and emotions. For example, people are partial towards their friends, peer group or race; emotions such as disgust or fear can interfere with moral deliberations, and so does our general mood; and the order in which people are asked to assess ethically relevant situations and the way that statistical information is presented make a difference to their judgments (see Sinnott-Armstrong 2011: 43–46). Even if those and many other factors do not apply to *all* of our ethical judgments, Sinnott-Armstrong claims that in their totality they provide evidence that a *significant percentage* of our moral judgments are distorted. The existence of these factors does not *prove* that the judgments which are affected by them are wrong, but the fact that they exercise an influence on moral deliberation undermines the trustworthiness of moral judgments in general. This is a problem for intuitionists insofar as foundationalism must claim that at least those judgments that are supposed to be non-inferentially justified are worthy of our trust.

In response, can't the intuitionist accept *with* Sinnott-Armstrong that we have reason to doubt many of our moral beliefs, but argue *against him* that the moral judgments intuitionism takes to be self-evident are the outcome of a reliable epistemic process (such as repeated careful deliberation involving mental maturity and the absence of strong emotions)? The strategy of this response is to divide the class of moral judgments into two subgroups: The first subgroup consists of all those moral judgments where psychological defeaters are prevalent and whose credence is cast into doubt, while a second and presumably much smaller subgroup contains privileged judgments that are the result of a reliable belief-forming process which excludes the distorting influences that characterize the members of the first subgroup. They therefore satisfy higher epistemic standards. Foundationalists could claim that self-evident beliefs belong to the second subgroup.

To block such a move, Sinnott-Armstrong introduces two epistemic principles. First, he argues that if a belief A falls into a group of beliefs B that possess a property C with a certain likelihood, our

default assumption should be that belief A has the same chance as any other belief in group B of having property C. If one wants to claim that belief A falls into another group D with a higher or lower chance of having property C, one needs provide evidence for that claim. By doing so, one makes use of a justificatory inferential structure as part of the overall justification (see Sinnott-Armstrong 2011: 48). This sounds plausible: If I happen to know that the vast majority of pocket watches show the wrong time and I encounter a random pocket watch, I should assume that it is inaccurate unless I can confirm that this particular one is more reliable than most other pocket watches. Sinnott-Armstrong's second epistemic principle is that if one has reason to conclude that a certain belief might with a certain likelihood be wrong, then one is not justified in maintaining it (see Sinnott-Armstrong 2011: 50).

This confronts the intuitionist with a dilemma. On its first horn, the moral judgments that are the source of self-evident beliefs are subject to psychological defeaters to the same extent as most other moral judgments (i.e. they belong to the first subgroup) and are therefore not trustworthy. Thus, we should suspend judgment regarding their truth. On the second horn, the intuitionist might claim that self-evident moral beliefs fall into the second subgroup and are therefore more reliable than other moral judgments. But it then follows from the first epistemic principle that such a claim stands in need of inferential justification and is therefore incompatible with foundationalism (see Sinnott-Armstrong 2011: 55–57).

The upshot is that given the empirical evidence that many of our moral intuitions have a significant chance of being false, there can be no non-inferentially justified moral belief; this undermines the possibility of a foundationalist moral epistemology and thereby questions the plausibility of one of the pillars of ethical intuitionism.

While some philosophers support Sinnott-Armstrong's attack on foundationalism,⁶ his argument has also provoked a number of critical responses. Here I will focus on what seems to me the most forceful argument against it to date.⁷ Variations on this argument have been offered in a number of places.⁸ In what follows, I will only discuss the version brought forward by Ballantyne and Thurow (2013), but I believe that my comments also apply to other formulations of the argument.

3 Ballantyne and Thurow's Argument against Sinnott-Armstrong

Nathan Ballantyne and Joshua Thurow concede to Sinnott-Armstrong that psychological factors undermine the trustworthiness of moral judgments in general. Against Sinnott-Armstrong, however, they make the case that the presence of these factors does not speak against the possibility of non-inferential justification. To this end, they introduce an example to help analyse the relationship in which the various beliefs that are relevant for the justification of a non-inferentially justified belief stand to one another. This allows them to argue that exclusion of distorting psychological factors can be accomplished without thereby adding to the justification of a belief and hence without undermining foundationalism. Here is their example ("case 1"):

⁶ His argument has, among others, been defended by Joyce (2006: 216–219) and Naddelhoff and Feltz (2008).

⁷ There is a popular strand of criticism that I will leave aside for now but will return to in the last section. It targets Sinnott-Armstrong's empirical assumption, making the case that the psychological defeaters as he presents them provide insufficient grounds for undermining the justification of non-inferentially held moral beliefs.

⁸ See Tolhurst (2008: 81f.), Tropman (2011: 364f.) and Ballantyne and Thurow (2013: 414). The argument has also been embraced by Stratton-Lake (2016).

McCoy visits the local widget factory and sees what seems to be a red widget being carried along a conveyor belt. He believes that the widget is red (call this belief "B"). Soon enough, a stranger approaches McCoy and says that the widgets are actually white but are illuminated by red lights. (Call this event "D".) Upon seeing this conversation, another stranger—who seems to McCoy to be a factory employee—tells McCoy not to listen to the other stranger: he is a trickster; McCoy is told, who likes to mess around with visitors. (Call this second event "F".) (Ballantyne and Thurow 2013: 413)

First, McCoy has a sensory experience which provides him with non-inferential justification for his belief B. His justification is, however, undermined by a defeater D which McCoy in turn calls into question by employing the defeater-defeater F. The notion of a defeater that is at play here is defined as follows:

R is a *defeater* for P as a *prima facie* reason for Q if and only if P is a reason for S to believe Q and R is logically consistent with P but (P & R) is not a reason for S to believe Q.⁹

In accordance with this definition, the stranger's objection (D) is consistent with McCoy's belief (B), but McCoy's original reason for believing B—his visual evidence—counts no longer in favor of B once the defeater is taken into account.

Based on this interpretation of the example, Ballantyne and Thurow argue that B is "supported by experiences or seemings alone" (Ballantyne and Thurow 2013: 414). The defeater-defeater only targets the defeater "without giving evidence for thinking B is true or that B was reliably formed". Therefore, F *preserves* the justification for B, but it does not *support* it. The factors that had justified B prior to the occurrence of the defeater continue to do so once the defeater-defeater is in place. This, Ballantyne and Thurow argue, makes room for the claim that the defeater-defeater is not part of a justificatory inferential structure supporting B.

Since Sinnott-Armstrong's criticism only targets the claim that the *justification* for our moral beliefs can be non-inferential, Ballantyne and Thurow argue that based on the foregoing distinction, there is no tension between the claim that the *explanation* that a given belief is not subject to a defeater is based on inferences and the claim that the *justification* itself is non-inferential. This is because the production and the preservation of justification are separate functions, and the fact that the latter is based on inferences does not imply that the former is as well.

4 An Objection to Ballantyne and Thurow

In order for Ballantyne and Thurow's argument to succeed, not only must their interpretation of case 1 be plausible, but their claim that defeater-defeaters are not part of the justification must generalize and extend to other cases, in particular to those involving moral judgments. I want to argue that this second assumption is false. While it seems correct that in the scenario introduced in the last section, the defeater-defeater falls outside the scope of the justification, there is an alternative explanation for

⁹ Pollock (1987: 484); see also Pollock and Cruz (1986: 37). Ballantyne and Thurow adopt Pollock's understanding of the term (See Ballantyne and Thurow (2013: 412f)). Note that Pollock uses the term "prima facie" in a way that differs from intuitionists like Ross. For Pollock, it means "defeasible" insofar as "further information can make us withdraw" support for the claim that P is a reason for Q. For Ross, a prima facie duty is a consideration that would be the duty all things considered (or 'duty proper'), if it was the only relevant consideration. In the presence of other considerations, it can be outweighed, even though it remains a reason. See Ross (1930: 19f).

why this is the case. This alternative explanation is sensitive to the kind of defeater that is brought forward against a certain belief, and in the cases of moral beliefs that Sinnott-Armstrong is interested in, it is an open question whether defeater-defeaters fall within or outside of the scope of the justification. If correct, this would serve as a rejoinder to Ballantyne and Thurow's argument.

To begin with, there are two features that can be ascribed to the defeater in case 1. First, the defeater does not originate from a reliable or authoritative source but rather comes from a "stranger" who is described as a "trickster [...] who likes to mess around with visitors".¹⁰ Second, the defeater as presented lacks intrinsic plausibility—absent further explanation, it seems implausible that the widgets would be illuminated just to fool the color perceptions of strangers. We can call defeaters of this sort "weak". In contrast to weak defeaters, *robust* defeaters (i) originate from a reliable and authoritative source and (ii) are intrinsically plausible.¹¹ While the defeaters that are relevant in the present context are, in both their weak and robust variants, aimed at undermining the reason in favor of the belief in question, they differ regarding their credibility.¹² A weak source and a lack of intrinsic plausibility detract from the credibility of a defeater and give less compelling reason to take it seriously, while defeaters that originate from a trustworthy source and are intrinsically plausible have the opposite effect.¹³ The following example (case 2) illustrates what counts as a robust defeater:

At night, Alice steps outside of her house, looks at the dark sky and believes that she sees Epsilon Indi, one of the smallest stars visible to the naked eye ("H"). But as her friend and long-time stargazer Stella reminds her, the sky above their town has been under a thick cloud of fog for years, due to a nearby coal-burning power plant. She adds that this makes it very unlikely that a star the size of Epsilon Indi could be detected ("J"). The following day, Alice reads in the newspaper that yesterday, the power plant had to be shut down due to an unexpected temporary malfunction and that consequently, the night sky had been clear ("L").

In contrast to McCoy and the trickster in case 1, Alice has reason to trust her well-informed friend Stella. Additionally, the defeater is plausible: neither the existence of the power plant nor the fact that coal-burning power plants cause foggy skies and are therefore bad for stargazing are in doubt.

If Alice wanted to justify to her neighbour her belief that she has seen *Epsilon Indi*, what would we expect her to include as part of her justification? —certainly, her visual experience, but if she were to stop there, this would not suffice to convince her interlocutor, who is also familiar with the

¹⁰ Ballantyne and Thurow's description downplays the credibility of the source of the defeater more than most typical descriptions of the same example. In Plantinga (2000: 359), it is the shop superintendent who brings forward the defeater, and in Evans and Smith (2013: 359), it is the factory owner. If my own account (see below) is correct, the way the defeater is characterized matters for the intuition Ballantyne and Thurow are trying to elicit, and it is not an accident that they characterize the source of the defeater in the way they do.

¹¹ These two criteria are not meant to be exhaustive, as I have a more limited goal: to demonstrate that two factors—the trustworthiness of the source and the intrinsic plausibility of the defeater—play a role (an important one) in determining how credible a defeater is. Other factors, such as the question of how many of my firmly held beliefs would be affected if the defeater were true or whether the defeater holds true in many nearby possible worlds, might have to be added for a complete account of the distinction between robust and weak defeaters.

¹² It is also possible for defeaters to be weak in the sense that they only slightly affect the likelihood of the reason in favor of the belief being true, rather than aiming at undermining it completely. As an example, in case 1, the defeater could state that only once per month or even only once per year, the factory is illuminated by red lights. This counts as a defeater, but as one that, if true, would not significantly undermine the justification in favor of seeing a red widget. Here, I am only concerned with defeaters that are not weak in terms of their content but rather aim at dealing a full blow to the reason in favor of the belief.

¹³ It is of course possible that some defeaters fall into the middle of the spectrum. Here, I only want to argue that there are clear cases of weak and robust defeaters.

smog caused by the local power plant but probably has not heard about the plant's temporary shut-down. For this reason, Alice should also mention the unusually favourable circumstances under which her observation took place.¹⁴ However, this stands in contrast to the intuition that case 1 elicits. What explains this difference?

Merely pointing towards the different kinds of defeaters in cases 1 and 2 does not yet *explain* why weak defeaters seem—at least in the cases under consideration—to lead to a narrow-scope justification, whereas robust defeaters result in wide-scope justifications. For this purpose, we must go beyond the defeaters and consider their function within the larger context of justification.

Being justified is a favorable evaluative status for a specific belief which consists of considerations that make it plausible that the belief in question is true. A justification fails in its purpose if it does not establish this claim. In the same vein, Laurence Bonjour argues that

if finding epistemically justified beliefs did not substantially increase the likelihood of finding true ones, then epistemic justification would be irrelevant to our main cognitive goal and of dubious worth. It is only if we have some reason for thinking that epistemic justification constitutes a path to truth that we as cognitive beings have any motive for preferring epistemically justified beliefs to epistemically unjustified ones (Bonjour 1985: 8).

Reasons in favor of the belief as well as considerations which ensure that those reasons are not undermined contribute towards the goal of demonstrating that a belief is likely to be true.¹⁵ In the presence of defeaters, neither the reasons that would in isolation have supported the belief nor the defeater-defeaters are taken by themselves to be capable of turning an unjustified belief into a justified belief. It is only their conjunction that achieves this purpose, and hence, *both elements should count as part of the justification*. However, this claim calls for a qualification: as it stands, it does not account for the different intuitions behind cases 1 and 2. The reason is that as long as defeaters with low credibility have not been discounted, all defeaters would have to be addressed with equal urgency since each of them would, if undefeated, significantly affect the likelihood of the belief being true.

This would lead to ramifications that extend beyond the problem of explaining our intuitions in a few isolated scenarios. Given the large number of defeaters that could be raised for each situation,

¹⁴ If, however, this is a fact that is well-known both to Alice and to her interlocuter, then there is no point in including the defeater-defeater as part of the justification.

¹⁵ Some philosophers such as William Alston have drawn a distinction between the concept of being justified in believing *p* (i.e. the state of being justified in a belief) and the concept of justifying *p* (i.e. the activity of justifying a belief), arguing that it is important to keep them apart for the following reason: "The crucial difference between them is that while to justify a belief is to marshal considerations in its support, in order for one to be justified in believing that *p* it is not necessary that one have done anything by way of an argument for *p* or for one's epistemic situation vis-à-vis *p*. Unless one is justified in many beliefs without arguing for them, there is precious little one can justifiably believe" (Alston 1991: 71). For intuitionists like Ross on the other hand, the state of being justified in believing *prima facie* duties seems tied to the process by which we come to believe them; it is doubtful whether agents who have not undergone this process would also count as justified. This process can take various forms, for example, careful reflection on the meaning of a *prima facie* duty or the observation of various instantiations of the breach of this duty ('intuitive induction'). From this, we come to the conclusion that part of the meaning of, say, promise breaking is that it is *prima facie* wrong (see Ross (1930: 34–39)). However, this does not lead to the consequence that there would be "preciously little one can justifiably believe": first, Ross' focus is not on the totality of beliefs; this leaves it open how justification works for other kinds of beliefs. Second, given the centrality of basic moral beliefs and their place in life, most competent agents will naturally go through the process of intuitive induction (and at least some will reflect on their normative commitments as they mature). Hence, even regarding the subgroup of basic moral beliefs, most agents will count as justified on Ross' theory. For the stronger claim that, in most cases, "being justified" should not be separated from the activity of justification, see Almeder (1999: 92). I thank two anonymous referees for raising this issue.

justifications which have to address all of them would be enormously large and complex, making the concept worthless for its practical purpose: It would be impossible for finite beings to justify most of their beliefs in this way.¹⁶ The qualification that is needed in order to address this problem limits the defeaters that need to be taken into account to those which would, in the eyes of what Quinn calls “intellectually sophisticated adults” (see Quinn 1985: 482), substantially affect the likelihood of the belief’s being true and hence reduce its credibility below the threshold required for justification. This excludes any highly dubious considerations which—despite the fact that they would, if true, undermine the evidence in favor of the belief—lack a sufficient degree of credibility. However one fills out the notion of intellectually sophisticated adults (or “mature adults” in Ross’ terminology), as long as they need to assess the credibility of objections to their belief, something akin to the distinction between weak and robust defeaters will be part of their mental toolset.

The position according to which weak defeaters fall outside the scope of justification while robust defeaters are part of it enjoys two advantages: it addresses the difference in intuition between cases 1 and 2—in the first case, the defeater is weak, while in the second, it is strong—and it does so while offering an explanation that is not ad hoc: Given the purpose of justification, only credible factors should be addressed.

For Sinnott-Armstrong’s argument to succeed, then, it must be the case that there are robust defeaters which (as long as they remain undefeated) undermine support for the justification of moral judgments. If the psychological defeaters are—as in Ballantyne and Thurow’s example—of the weak kind, they will play no role in the justification of moral beliefs, and foundationalists will win the day. But whether or not this is so is a substantive question on its own that has to be solved by looking at the content and the source of those defeaters. It cannot be argued from the start—as Ballantyne and Thurow suggest—that for formal reasons, defeaters for psychological defeaters are always and necessarily excluded from the scope of the justification for moral beliefs. If Sinnott-Armstrong’s argument fails (and below I will argue that it does), this will be for different reasons.

5 Why Defeater-Defeaters Pose no Problem for Foundationalism, Even if they Are Part of the Justification for Moral Beliefs

Now that we have reason to assume that defeater-defeaters can be part of the justification for moral beliefs, friends of foundationalism have two main options.¹⁷ First, they can argue that

¹⁶ To count as a defeater, there is no need for it to actually be raised for each situation where it applies. It is sufficient that an intellectually sophisticated adult should, upon reflection, be aware that it could be raised. This includes all sorts of conspiracy theories, far-fetched skeptical scenarios and countless other defeaters.

¹⁷ There are other ways to criticize Sinnott-Armstrong’s argument which I will, however, leave aside. For example, one could distinguish between concrete beliefs about particular situations and general beliefs about principles and make the case that only concrete beliefs are susceptible to psychological defeaters. This would leave general beliefs such as the *prima facie* duty of benevolence untamished by Sinnott-Armstrong’s attack. Since the category into which a moral belief falls is clear from the start, there is no need for inferential confirmation. This line of argument can be traced to Ross, who argues that we can ascribe certainty only to abstract principles and that our judgments in concrete cases are “more or less probable opinions” (1930: 31.). A similar move is described by MacIntyre when he discusses how Calvinists in the seventeenth century were worried that fallible human beings could not know with certainty that God exists. The solution was to distinguish between concrete judgments regarding everyday objects and a privileged group of judgments about abstract matters. In virtue of being abstract, the latter group of judgments was considered less fallible (MacIntyre 1988: 241–259). I thank [...] for raising this issue. Yet another reply to Sinnott-Armstrong can be found in van Rooijen (2014), who argues that some moral beliefs could be justified in virtue of their non-inferential plausibility in addition to coherentist considerations.

the empirical evidence Sinnott-Armstrong provides is insufficient and does not undermine the trustworthiness of our moral intuitions. A number of critics have argued along these lines,¹⁸ and while it is certainly a legitimate way to question Sinnott-Armstrong's case, there remains a serious drawback to this strategy: Even if successful, arguments of this kind are vulnerable to new and more conclusive studies which provide better evidence.¹⁹ Given that all that skeptics of the reliability of moral judgments need is a cumulative case which shows that *in their totality*, psychological defeaters of various sorts cast doubt on the trustworthiness of moral intuitions, one might worry that, over time, a still more substantive body of data will emerge and make that reply seem less and less plausible. It would therefore be unwise for foundationalists to place all their bets on this first strategy. In what follows, I will therefore assume that the psychological defeaters Sinnott-Armstrong presents are of the robust kind and undermine, if no defeater-defeaters can be found against them, our trust in moral judgments.

I want to pursue a second option in defense of foundationalism. It accepts not only that the psychological defeaters Sinnott-Armstrong presents need to be addressed before foundationalist moral beliefs can be justified, but also goes further than Ballantyne and Thurow in granting that defeater-defeaters provide inferential support in favor of the moral beliefs in question. Despite these concessions, I believe that foundationalism in moral epistemology remains a viable position that solves the regress-problem and avoids coherentist forms of justification or the reduction of moral to non-moral beliefs. I will address these challenges in turn.

The immediate worry this position raises is that it fails to address the main problem raised by Sinnott-Armstrong: If foundationalists employ defeater-defeaters in order to undermine psychological defeaters, the justification involves inferential processes. This seems to correspond to the problem brought up by the second horn of Sinnott-Armstrong's dilemma.²⁰ However, there are two ways in which inferential processes could be taken to be problematic for foundationalism. First, it could be said that according to foundationalism, the justification base must not consist of beliefs which provide *inferential support* for foundational moral beliefs. Merely self-evident beliefs which, once they have been properly understood, provide immediate evidence for their content, may be part of the justification for foundational moral beliefs. Some of what Sinnott-Armstrong says can be taken to lend support to such a position, for example when he defines his target as the view that "the believer is justified regardless of whether the believer is committed to any justificatory inferential structure" (Sinnott-Armstrong (2011: 40; see also 38). A second way to interpret foundationalism in moral epistemology is that no element that is part of the justification base must stand itself in need of inferential justification. Here, the focus is on the epistemic status of the elements that are part of the justification base and not on their function regarding the justification of the belief in question. At some places, it seems that it is this second view that Sinnott-Armstrong has in mind, for example when he argues that the kind of non-inferentiality that ethical intuitionism "requires that the believer would still be justified even if the believer did not have as much as a disposition to accept any propositions that are, entail, or support any propositions that provide epistemic support for what is believed" (40; see also 41 and his Sinnott-Armstrong 2006b: 184–191). The concern that is expressed here is not with the idea that there might be various

¹⁸ Demaree-Cotton (2016), Smith (2010), Tolhurst (2008), and Shafer-Landau (2008) (even though this is not the main thrust of Shafer-Landau's argument).

¹⁹ See, for example, Naddelhoff and Feltz (2008), who identify additional cognitive biases to which moral judgments are subject and who argue that this strengthens Sinnott-Armstrong's case.

²⁰ I thank an anonymous referee for pressing me to clarify this issue.

kinds of considerations which support a certain belief, but with the question of whether these considerations stand themselves in need of justification. What makes the second interpretation look more plausible is the overall question foundationalists are trying to address. This becomes clear when Sinnott-Armstrong argues that views which deny the claim just quoted are “not enough stop the skeptical regress, opposition to which motivates many moral intuitionists” (40). If the second interpretation of foundationalism is correct, the second horn of Sinnott-Armstrong’s dilemma has to be read in light of it. It hence raises the worry that the employment of defeater-defeaters in response to psychological defeaters contradicts foundationalism insofar as those defeater-defeaters stand themselves in need of justification and fail to put a stop to the regress.

The question is whether the proposed solution – namely, that defeater-defeaters provide inferential support for foundational moral beliefs – is vulnerable to this objection. At first sight, this seems to be the case. After all, beliefs regarding defeater-defeaters (e.g. the belief that the process of reflection has not been subject to emotional distortions or partiality) are empirical and stand in need of further justification. It therefore appears that the thread of a regress has not been averted. However, ethical intuitionists can reply that the challenge they are addressing is limited to a skepticism regarding the justification of certain *moral* beliefs. There is no need for them to take on more philosophical baggage than necessary and to address issues outside their scope such as the justification of *non-moral* beliefs. While defending the possibility of non-inferential justification for moral beliefs, proponents of foundationalism are not even committed to an *agnostic* stance regarding the justification of belief in empirical facts; instead, they should simply be allowed to take for granted that other kinds of true beliefs can be justified in some non-problematic way.

This view entails nothing that classical intuitionists or more recent proponents of ethical intuitionism should find difficult to accept. Ross believes that we only come to know self-evident truths as the result of a cumulative process of deliberation that stretches over many generations (1930: 41). In order to assess which of the ideas that others have come up with are worth pursuing and which would better be disregarded, we would have to identify biases, prejudices and other sorts of bad thinking, and this involves non-moral premises. In Audi’s definition of self-evidence quoted at the beginning, it is “in virtue of understanding” that an agent gains justification. Audi explicitly rules out “clouded” or “distorted” forms of understanding (Audi 2004: 49 and Audi 1999: 207). This includes not only cognitive shortcomings such as a mistaken understanding of what is logically implied by a proposition, but also “biased background beliefs” (Audi 2008: 16). The requirement that agents need, as part of their justification for self-evident beliefs, to exclude psychological defeaters therefore also does not stand in tension with Audi’s views on foundationalism.

The assumption that the justifiability of non-moral premises has to be taken for granted is also compatible with Sinnott-Armstrong’s challenge to intuitionism. Not only that, but while adopting a skeptical standpoint regarding moral beliefs, Sinnott-Armstrong even presupposes that propositions about empirical facts can be justified. After all, if the justification for true empirical claims leads to an epistemic regress, this would also affect the premises regarding the psychological defeaters which Sinnott-Armstrong’s argument is built upon. Hence, proponents of his challenge to intuitionism need to reject skepticism regarding the justifiability of empirical beliefs. But if empirical beliefs raise no epistemological difficulties, shouldn’t foundationalists by extension also be allowed to employ empirical assumptions to exclude the presence of psychological defeaters? As just argued, proponents of Sinnott-Armstrong’s challenge are in no position to make the case that employing those assumptions leads to an

epistemic regress. Or, perhaps, does the use of defeater-defeaters instead invite the charge that foundationalists are deriving moral from non-moral beliefs?

In response to this worry, the function of the defeater-defeaters that are relevant in the present context is to enable considerations that speak in favor of a moral belief—e.g. that lying is *prima facie* wrong—to retain their reason-giving force despite the presence of defeaters which would otherwise undermine them. Hence, no normative conclusions regarding morality can be derived from the defeater-defeaters.²¹ As I have shown above, this does not contradict the claim that they belong to an inferential structure which as a whole justifies the normative belief in question.

Consequently, intuitionists can help themselves to defeater-defeaters while (a) upholding the claim that some moral beliefs are non-inferentially justified and (b) avoiding the charge of an epistemic regress or derivation of moral from non-moral beliefs.

Once intuitionists are licensed to use defeater-defeaters to strengthen the credibility of moral intuitions, they are in a position to demonstrate that the defeaters Sinnott-Armstrong introduces do not apply to all moral beliefs; cases where defeater-defeaters are available might, for example, include those found on Ross's list of *prima facie* duties. In fact, the limitations that Ross imposes on candidates for self-evident moral beliefs such as mature agency and careful reflection (see Ross 1939: 21) can be read as umbrella terms that encompass all defeater-defeaters which are necessary to deal with the kind of psychological defeaters that Sinnott-Armstrong discusses.²²

This casts into doubt Sinnott-Armstrong's argument that intuitionism is incompatible with an inferential structure. Non-normative premises which are not used to derive normative beliefs but which establish the absence of psychological defeaters do play a crucial role in the process of justification, but they do not undermine the status of justification as non-inferential in the sense intended by the foundationalist: No regress occurs since the foundationalist is entitled to assume that non-normative premises can be justified. The burden of showing how this can be done, however, rests on somebody else's shoulders.

Acknowledgments For helpful discussions and comments on earlier drafts of the paper, I would like to thank Bernd Goebel, Christoph Halbig, Josh May, Sebastian Muders, Stefan Riedener, Eva Schmidt and Felix Timmermann.

References

- Almeder R (1999) Harmless naturalism. The limits of science and the nature of philosophy. Open Court, Chicago and La Salle
- Alston WP (1991) Perceiving god. The epistemology of religious experience. Cornell University Press, Ithaca
- Audi R (1999) Self-evidence. *Philos Perspect* 13:205–228
- Audi R (2004) The good in the right: a theory of intuition and intrinsic value. Princeton University Press, Princeton

²¹ At the same time, it is the case that epistemic duties can be derived from defeater-defeaters: if it is correct, it follows among other things that I should no longer believe that the defeater undermines the original belief.

²² The problem that foundationalist (or “basic”) beliefs seem to be undermined by defeaters is also discussed by Alvin Plantinga in connection with religious claims. Similar to Ross, Plantinga does not believe that a rejection of those defeaters that has an inferential structure undermines the claim that those beliefs are “not accepted on the basis of other propositions” (Plantinga 2000: 174). For his discussion of defeaters, see Plantinga (2000: part iv).

- Audi R (2008) Intuition, inference, and rational disagreement in ethics. *Ethical Theory Moral Pract* 11(5):475–492
- Ballantyne N, Thurrow JC (2013) Moral intuitionism defeated? *Am Philos Q* 50(4):411–421
- Bonjour L (1985) *The structure of empirical knowledge*. Harvard University Press, Cambridge
- Demaree-Cotton (2016) Do framing effects make moral intuitions unreliable? *Philos Psychol* 29(1):1–22
- Evans I, Smith ND (2013) *Knowledge*. Policy Press, Cambridge
- Huemer M (2005) *Ethical Intuitionism*. Palgrave Macmillan, Hampshire.
- Joyce R (2006) *The evolution of morality*. MIT Press, Cambridge
- MacIntyre A (1988) *Whose justice? Which rationality?* University of Notre Dame Press, Notre Dame
- Naddelhoff T, Feltz A (2008) The actor–observer bias and moral intuitions: adding fuel to Sinnott-Armstrong’s fire. *Neuroethics* 1:133–144
- Plantinga A (2000) *Warranted Christian belief*. Oxford University Press, New York
- Pollock JL (1987) Defeasible reasoning. *Cogn Sci* 11:481–518
- Pollock JL, Cruz J (1986) *Contemporary theories of knowledge*, 2nd edn. Rowman & Littlefield, Lanham
- Quinn PL (1985) In search of the foundations of theism. *Faith Phil* 2(4):469–486
- Ross WD (1930/1967) *The right and the good*. Clarendon Press, Oxford
- Ross WD (1939) *The foundations of ethics*. Clarendon Press, Oxford
- Shafer-Landau R (2003) *Moral realism: a defence*. Oxford University Press, Oxford
- Shafer-Landau R (2008) Defending ethical intuitionism. In: Sinnott-Armstrong W (ed) *Moral psychology, Vol II, The cognitive science of morality: intuition and diversity*. MIT Press, Cambridge, pp 83–96
- Sinnott-Armstrong W (2002) Moral relativity and intuitionism. *Philosophical Issues* 12:205–328
- Sinnott-Armstrong W (2006a) Moral intuitionism meets empirical psychology. In: Horgan T, Timmons M (eds) *Metaethics after Moore*. Oxford University Press, Oxford, pp 339–365
- Sinnott-Armstrong W (2006b) *Moral skepticisms*. Oxford University Press, Oxford
- Sinnott-Armstrong W (2008) Framing moral intuitions. In: Sinnott-Armstrong W (ed) *Moral psychology, Vol II, The cognitive science of morality: intuition and diversity*. MIT Press, Cambridge, pp 47–76
- Sinnott-Armstrong W (2011) An empirical challenge to moral intuitionism. In: the new intuitionism. Graper Hernandez J (ed.) *Continuum*, London, pp 11–28
- Smith J (2010) On Sinnott-Armstrong’s case against moral intuitionism. *Ethical Theory Moral Pract* 13(1):75–88
- Stratton-Lake P (2002) Introduction. In: Stratton-Lake P (ed) *Ethical intuitionism: re-evaluations*. Oxford University Press, Oxford
- Stratton-Lake P (2016) Intuitionism in ethics. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy*, Winter edn. <https://plato.stanford.edu/archives/win2016/entries/intuitionism-ethics/>. Accessed 3 April 2018
- Tolhurst W (2008) Moral intuitions framed. In: Sinnott-Armstrong W (ed) *Moral psychology, Vol II, the cognitive science of morality: intuition and diversity*. MIT Press, Cambridge, pp 77–82
- Tropman E (2011) Non-inferential moral knowledge. *Acta Analytica* 26(4):355–366
- van Roojen M (2014) Moral intuitionism, experiments, and skeptical arguments. In: Booth AR, Rowbottom DP (eds) *Intuitions*. Oxford University Press, Oxford, pp 148–164

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Philipp Schwind¹

¹ Institute of Philosophy, Center for Ethics, University of Zurich, Zollikerstrasse 117, 8008 Zurich, Switzerland